

Accretion by the Galaxy

James Binney^{1,a} and Filippo Fraternali^{2,3}

¹ University of Oxford, Rudolf-Peierls Centre for Theoretical Physics, Keble Road, Oxford OX1 3NP, UK

² University of Bologna, Department of Astronomy, via Ranzani 1, 40126, Bologna, Italy

³ Kapteyn Astronomical Institute, Postbus 800, 9700 AV, Groningen, NL

Abstract. Cosmology requires at least half of the baryons in the Universe to be in the intergalactic medium, much of which is believed to form hot coronae around galaxies. Star-forming galaxies must be accreting from their coronae. H I observations of external galaxies show that they have H I halos associated with star formation. These halos are naturally modelled as ensembles of clouds driven up by supernova bubbles. These models can fit the data successfully only if clouds exchange mass and momentum with the corona. As a cloud orbits, it is ablated and forms a turbulent wake where cold high-metallicity gas mixes with hot coronal gas causing the prompt cooling of the latter. As a consequence the total mass of H I increases. This model has recently been used to model the Leiden-Argentina-Bonn survey of Galactic H I. The values of the model's parameters that are required to model NGC 891, NGC 2403 and our Galaxy show a remarkable degree of consistency, despite the very different natures of the two external galaxies and the dramatic difference in the nature of the data for our Galaxy and the external galaxies. The parameter values are also consistent with hydrodynamical simulations of the ablation of individual clouds. The model predicts that a galaxy that loses its cool-gas disc for instance through a major merger cannot reform it from its corona; it can return to steady star formation only if it can capture a large body of cool gas, for example by accreting a gas-rich dwarf. Thus the model explains how major mergers can make galaxies “red and dead.”

1 Introduction

The Milky Way is typical of the galaxies that now dominate the cosmic star-formation rate (SFR): its luminosity lies extremely close to the characteristic luminosity L^* of the Schechter galaxy luminosity function, so it is one of the most massive galaxies that are still actively forming stars. The colour-magnitude diagram of the local stars implies that the SFR in the thin disc, which is the Galaxy's dominant component, has declined only by a factor of a few over the last 10 Gyr (Hernandez et al., 2000; Aumer & Binney, 2009). Since the SFR must depend strongly on the cold-gas content of the disc, this finding indicates that despite turning $\sim 5 \times 10^{10} M_{\odot}$ of gas into stars, the disc has only mildly depleted its cold-gas content, which is significantly less than $10^{10} M_{\odot}$. It follows that the Galaxy must be somehow replenishing its supply of cold gas.

Here we present a model of how the Galaxy acquires cold gas. This model was developed to explain observations of the H I content of nearby galaxies and has been recently applied to the qualitatively different data for our Galaxy. Although the model is still rather crude, and a number of aspects need to be worked out in greater detail, it ties together disparate data in a remarkably coherent manner, which suggests that the underlying physical picture is correct.

^a e-mail: binney@thphys.ox.ac.uk

2 Cosmological context

2.1 Missing baryons

Within the context of standard cosmological theory we can estimate the cosmic density of baryons in two distinct ways (e.g. Peacock, 1999): (i) by combining cosmic nucleosynthesis with measurements of the abundances of helium, deuterium and lithium, and (ii) from the power spectrum of fluctuations in the temperature of the cosmic microwave background. Method (i) yields the number of CMB photons per baryon, while method (ii) yields the fractions of the closure density contributed by baryons and by matter as a whole. Both methods point to baryons contributing $\sim 4\%$ of closure density and method (ii) implies that baryons contribute 17% of the matter density.

The gravitating mass of the Local Group can be reliably estimated from either the classical Kahn-Woltjer timing argument (e.g. Binney & Tremaine, 2008) or from a new formulation of it (Li & White, 2008), so we know that $\sim 8 \times 10^{11} M_{\odot}$ of baryons were originally associated with the Local Group. Observations of the stars and gas in the Local-Group galaxies give a total baryon content of $< 1.5 \times 10^{11} M_{\odot}$, less than a fifth of the expected amount.

It is generally believed that the missing baryons are contained in hot, diffuse gas. Unless this gas has escaped from the gravitational field of the Galaxy and the Local Group, its temperature must lie in a narrow range around the virial temperature T_{vir} of the confining gravitational well – at temperatures slightly smaller than T_{vir} , the gas is so narrowly confined to the centre of the well that its pressure near the Sun would exceed the pressure within clouds of well-studied cool gas. Gas with $T \simeq T_{\text{vir}}$ will be referred to as “coronal gas”.

In a rich cluster of galaxies the gravitational field is strong enough to compress coronal gas to densities at which it can be readily detected either through its X-ray emission (Gursky et al., 1971) or through distortion of the spectrum of the CMB via the SZ effect (Sunyaev & Zeldovich, 1972). In massive galaxy groups coronal gas can be detected through its X-ray emission, but in low-mass groups such as the Local Group the density and temperature of coronal gas is expected to be too low for detection in either X-rays or the SZ effect to be possible. However, we do have indirect indications of the existence of coronal gas within the Local Group. One indication is provided by ultraviolet observations of objects at high Galactic latitude, which reveal absorption by highly ionised species such as OVI along lines of sight that pass close to HI clouds (Sembach et al., 2003); models of the interface between coronal gas and HI predict the existence of regions rich in OVI. In fact, long before the emergence of modern cosmology Spitzer (1956) inferred the existence of coronal gas from the existence of absorption lines of interstellar Na and Ca in the spectra of high-latitude stars by arguing that without a confining medium at $\sim T_{\text{vir}}$ the clouds responsible for this absorption would quickly dissipate.

Another pointer to coronal gas is the gross asymmetry between the trailing and leading Magellanic Streams – the Stream is almost certainly formed of gas stripped from the SMC-LMC system. If clouds within the Stream were orbiting freely, the leading stream would be as extensive as the trailing stream. In fact it is much shorter than the trailing stream (Putman et al., 1998), a phenomenon that arises naturally if clouds of HI experience hydrodynamical drag as they move through ambient gas of density $\sim 10^{-4} \text{ cm}^{-3}$ (Mastropietro et al., 2009; Besla et al., 2010). The morphology of compact high-velocity HI clouds provides a similar argument for the existence of a low-density ambient medium: when well resolved, these clouds generally have a tadpole-like head-tail morphology indicative of distortion by motion through an ambient medium (Westmeier et al., 2005). Thus many lines of argument suggest that the missing baryons are contained in gas at $\sim 2 \times 10^6 \text{ K}$ that fills much of the Local Group.

2.2 Evidence for accretion

The history of a galaxy’s star formation is encoded in the colour-magnitude diagram of its stars, which can be probed either directly when the galaxy is close enough to resolve its stars, or indirectly through the Galaxy’s colours. The conclusion of many studies of CMDs and galaxy colours is that in spiral galaxies like the Milky Way, the SFR is only slowly declining. Only $\sim 10\%$ of the masses of the discs

of such galaxies is contributed by detected gas, so much more gas has been turned into stars than the discs currently possess. Hence either star-formation is about to cease in these systems, or they are constantly replenishing their gas supply. Since the SFR of a disc must depend on how much cold gas it has, the slowness of the decline in the SFR argues against a rapid drop in the stock of cold gas, and thus in favour of continuous replenishment. This conclusion is strongly reinforced by models of the chemical enrichment of our own disc, which require constant accretion of relatively metal-poor gas to explain why there are not more metal-poor stars near the Sun than are observed (the “G-dwarf problem” Pagel & Patchett, 1975).

2.3 Why isn't the Galaxy red and dead

To understand the impact of the hot IGM on disc galaxies it is natural to turn to rich clusters of galaxies, where the IGM can be directly observed. The cooling time of the IGM, being inversely proportional to the density, is shortest at the centre of the system, so this is where we expect to find the coolest gas. X-ray observations bear out this expectation, but with a surprising twist: although near the centre the cooling time is often much shorter than the Hubble time, there is very little gas at temperatures lower than $\sim T_{\text{vir}}/3$ (Peterson et al, 2003). Indeed, there is an almost complete absence of spectral lines from species such as Fe xvii that should form as gas cools through $\sim 10^6$ K. Thus near the centres of rich clusters, gas is radiating, but not cooling. It follows that its radiative losses are being offset by a heat source, and the obvious mechanism is mechanical heating by the bipolar outflows that are known to accompany accretion onto compact objects, in this case the black holes that reside at the centres of galaxies. In fact, mechanical heating by black holes was predicted by Binney & Tabor (1995) before observations revealed the absence of cool gas. Unlike alternative heating mechanisms (Ciotti & Ostriker, 2011), it is naturally self-regulating (Omma & Binney, 2004; Binney, 2005). Regardless of what stops the temperature of gas in clusters falling below $T_{\text{vir}}/3$, this phenomenon poses a puzzle in the context of disc galaxies because it suggests that the black holes at the centres of spiral galaxies should be able to prevent coronal gas cooling effectively, just as the black holes at the centres of clusters do. Indeed, the weakness of the X-ray emission around spiral galaxies compared to that from rich clusters bears witness to the relative ineffectiveness of radiative cooling in the coronal gas of spiral galaxies compared to that in rich clusters. Hence if mechanical heating is effective in rich clusters, why would it fail around disc galaxies?

A suggestion that the black hole at the centre of the Galaxy *does* effectively heat the coronal gas is provided by the fact that the bulk of the star formation in the Galaxy occurs not near the Galactic centre, where the pressure and density of the coronal gas must be largest and therefore its cooling time must be shortest. Instead the bulk of star formation occurs several kiloparsecs away from the centre, in the disc. Moreover, what star formation does take place near the Galactic centre, in the nuclear molecular disc, is likely fed by gas that has been driven in by the Galactic bar, from its corotation resonance to its inner Lindblad resonance. Thus we need to explain why galaxies at the centres of rich clusters long ago ran out of cold gas despite being enveloped in enormous quantities of strongly radiating coronal gas, while spiral galaxies manage to stay youthful and blue by replenishing their stocks of cold gas from their diffuse coronal gas. In particular we have to explain why in spiral galaxies cooling coronal gas falls onto the disc several kiloparsecs away from where the cooling time of the coronal gas is shortest.

3 Extraplanar HI

Sensitive studies of nearby spiral galaxies in the 21-cm line of H I reveal that significant quantities of H I lie more than a kiloparsec above or below the equatorial planes of many galaxies (Sancisi et al., 2008, and references therein). In edge-on galaxies such as NGC 891 this conclusion follows rather directly from the projected distribution of H I emission on the sky (although kinematic information is required to rule out the possibility that the material seen above and below the plane lies in an extended warped disc). In galaxies that are not seen edge-on, the existence of gas above and below the plane must be established by detailed kinematic modelling (Boomsma et al., 2008). H I that lies more than a kiloparsec from the equatorial plane is called “extraplanar H I” and constitutes an “H I halo”.

Studies of extraplanar H I reveal three indications that extraplanar H I is associated with star formation: (i) the radial extent of H I is similar to that of significant star formation; (ii) the mass of the extraplanar H I is correlated with the galaxy's total SFR; (iii) the galaxy's main H I layer reveals small holes, and these holes are often associated with not only the signs of recent star formation, but also nearby H I emission at anomalous velocities. The conclusion is inescapable that localised bursts of star formation in the disc give rise to supernova bubbles that blast nearby H I out of the disc and up into the H I halo.

From kinematic modelling of H I halos one can show that the azimuthal velocity of the H I drops quite steeply with distance from the plane (Oosterloo et al., 2007). This fact proves to be of considerable importance.

4 A model of extragalactic fountains

The idea that star formation will drive a circulation of gas in galaxies goes back a long way and has been explored by many authors (Bregman, 1980; Collins et al., 2002). Fraternali & Binney (2006, hereafter FB06) modelled the H I datacubes of NGC 891 and NGC 2403 as emission from an ensemble of clouds moving on ballistic trajectories. Their clouds are launched from platforms in the disc that are on circular orbits with a radial distribution that reflects the distribution of star formation. The clouds' velocities are of order h_v and have an angular distribution that has to be sharply concentrated around the Galactic poles. They found that with $h_v \sim 75 \text{ km s}^{-1}$ this model was able to account for most aspects of the data for both galaxies, despite the galaxies' strongly contrasting masses and inclinations. However, the fit to the data was deficient in two respects: the azimuthal streaming velocity of the H I halo was predicted to decline too slowly with distance from plane, and the data indicated a stronger bias towards inflow than the models could reproduce, even when clouds were assumed to be ionised and therefore invisible as they were shot upwards. Fraternali & Binney (2008, hereafter FB08) showed that these deficiencies were removed if the orbits and masses of clouds were modified by interaction with coronal gas. The successful models assumed that clouds sweep up the coronal gas they encounter on their paths, the rate of accretion being quantified by the e-folding time α^{-1} of a cloud's mass. Because the corona was assumed to be non-rotating, the accretion of this gas slows the azimuthal motion of clouds, as the data require. The accretion has two other beneficial effects: it causes infall to predominate over outflow as the data require, and it prevents net transfer of angular momentum to the corona by ram pressure. The second point is important because the moment of inertia of the part of the corona near the optical galaxy is small, and if not counteracted by accretion, ram pressure would quickly set this part of the corona rotating nearly as fast as the disc. Then the action of the disc on the corona would be that of a centrifugal pump. The values $h_v = 75 \text{ km s}^{-1}$, $\alpha = 1.5 \text{ Gyr}^{-1}$ provided good fits to the data for both galaxies. The models predicted net accretion rates of 2.9 and $0.8 \text{ M}_\odot \text{ yr}^{-1}$ for NGC 891 and NGC 2403, respectively, quite similar to the galaxies' star-formation rates.

Marinacci et al. (2010) illuminated the results of FB08 by simulating the motion of an H I cloud through coronal gas. They found that the cloud was inevitably ablated by the hot gas flowing over its surface. However, if the density of the coronal gas and the metallicity of the cloud were high enough, the total mass of cool gas, visible as H I, would nonetheless increase because in the cloud's turbulent wake ablated cool gas would cause coronal gas to cool and become neutral. Consequently, they argued that the model of FB08 was fundamentally correct providing their "clouds" were understood to be clouds plus wakes of cool gas. They pointed out that in external galaxies H I observations lack the sensitivity to detect an individual stream of the type they predicted, but the H I datacube of our Galaxy does contain suitably elongated features. Marinacci et al. (2011) presented more elaborate simulations of clouds moving through coronal gas and showed that there *is* a net transfer of momentum from the cloud to the corona if their relative velocity exceeds a threshold $\sim 50\text{--}85 \text{ km s}^{-1}$ that increases with the coronal density. The existence of this threshold velocity suggests that coronae spin at a non-negligible rate, but slower than the disc by $80\text{--}120 \text{ km s}^{-1}$.

5 Modelling the LAB datacube

Marasco et al. (2011) applied the model of FB08, modified to include the insights gained by Marinacci et al. (2010, 2011), to the datacube from the Leiden-Argentine-Bonn (LAB) all-sky survey of Galactic H I (Kalberla et al., 2005). From these data, which are dramatically different in type from those from which the model was developed, they re-determined the model’s parameters, h_v and α . They found $h_v = 70 \text{ km s}^{-1}$ in line with the values required by external galaxies, and $\alpha = 6.3 \text{ Gyr}^{-1}$, a factor ~ 4 larger than that favoured by external galaxies. Most of this difference is attributable to the adoption by FB08 of a non-rotating corona: when the Galaxy’s corona is assumed to be non-rotating, the LAB data require $\alpha = 2.5 \text{ Gyr}^{-1}$. However, for a rigorous comparison of results for our galaxy and external galaxies the current model needs to be fitted to the data for external galaxies.

The key differences between the model fitted to the LAB data and the model used by FB08 are (i) the adoption of a rotating corona, and (ii) a more sophisticated parametrisation of how gas is ionised as it leaves the disc and neutral on its return: gas becomes neutral when its vertical velocity is smaller than its value on ejection by a factor $1 - f_{\text{ion}}$. The value of f_{ion} is not tightly constrained by the data, but values $f_{\text{ion}} \simeq 0.3$ work well, so gas becomes neutral quite soon after it is ejected.

From the model that fits the LAB data we can read off the global properties of the Galaxy’s H I layer. The rotation speed falls $\sim 20 \text{ km s}^{-1}$ below the circular speed 1 kpc above the plane and then by a further $\sim 10 \text{ km s}^{-1}$ with each kiloparsec. A similar profile was obtained for NGC 891 by Oosterloo et al. (2007). The scaleheight of the H I layer increases strongly with Galactocentric distance R from $\sim 0.35 \text{ kpc}$ at $R = 4 \text{ kpc}$ to 1.75 kpc at $R = 12 \text{ kpc}$. The rate of accretion of pristine gas rises from near zero at $R < 3 \text{ kpc}$ to a peak value $\sim 7 \times 10^{-3} \text{ M}_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ at $R = 9 \text{ kpc}$ and then falls to zero at $R > 13 \text{ kpc}$. Overall the Galaxy is predicted to accrete $1.6 \text{ M}_{\odot} \text{ yr}^{-1}$ of H I or $2.3 \text{ M}_{\odot} \text{ yr}^{-1}$ of gas when helium is included. For comparison, the Galaxy’s star-formation rate is $\sim 3 \text{ M}_{\odot} \text{ yr}^{-1}$ (Diehl et al., 2006), of which $\sim 1 \text{ M}_{\odot} \text{ yr}^{-1}$ will be provided by dying stars. In addition to gas accreted through the modelled fountain clouds, Sancisi et al. (2008) estimate that accretion of high-velocity clouds, which are not accounted for by the model, provides $0.2 \text{ M}_{\odot} \text{ yr}^{-1}$ of gas.

The parameter α has the job of quantifying the rate at which the mass of cold gas associated with an ejected cloud increases over time, and by conservation of momentum it predicts the rate of decrease of the velocity centroid of the cold gas. The detailed hydrodynamical simulations of Marinacci et al. (2010, 2011) predict these same quantities from ab-initio physics. Remarkably the value of α required to fit the LAB data provides excellent fits to plots of mass and velocity versus time from the simulations. This result makes sense only if the model has correctly captured the basic physics of the Galactic fountain.

6 Conclusions

Star-forming galaxies such as our own must accrete intergalactic gas. Our current understanding of why cool-core galaxy groups and clusters are deficient in star formation makes it paradoxical that the only galaxies able to accrete are those hosted by less deep gravitational potential wells in which the coronal gas radiates too weakly to be detected. It is also strange that in these galaxies accretion does not occur at the centre, where the pressure must be highest and the cooling time shortest, but kiloparsecs away in a disc. A promising resolution of this paradox is that galactic fountains reach up and grab coronal gas, bringing it down into the star-forming disc. The key to extracting gas from the corona far from the centre is the provision of “seed” cool gas, which lowers the cooling time of the coronal gas with which it mixes in the wake of each fountain cloud.

If this model is correct, galaxies which lose their cool-gas discs, for example in a major merger, will not be able to reform them from coronal gas; their only hope of continued star formation is accreting a substantial body of cool gas, for example from a gas-rich dwarf. A galaxy that does not have a cool-gas disc will not drive a fountain, and without a fountain gas can be cooled out of the corona only at the centre, where it will feed the central black hole and stimulate it into an outburst that will reheat the corona and prevent further cooling.

Observations of the H I halos of external galaxies seem to require interactions between fountain clouds and coronal gas similar to those predicted by basic physics: clouds must be exchanging mass

and momentum with the corona. The data require that the clouds gain mass, while hydrodynamic simulations clearly predict that clouds are steadily ablated by the corona. However, for an astronomically plausible range of parameters, condensation of coronal gas in the turbulent wake of each cloud yields the predicted changes in the total mass and mean velocity of the cool gas. The rate of accretion predicted in this way is in excellent agreement with that required to sustain star formation, and there is remarkable agreement between the rates of mass and momentum exchange required by the observations and hydrodynamical simulations.

In summary, it seems likely that supernova-driven fountains provide the dominant mechanism through which disc galaxies collect gas from intergalactic space, at least at redshifts $z \lesssim 1$.

References

- Aumer M., Binney J., 2009, MNRAS, 397, 1286
 Besla G., Kallivayalil N., Hernquist L., van der Marel R.P., Cox T.J., Keres D., 2010, ApJ, 721, 97
 Boomsma R., Oosterloo T. A., Fraternali F., van der Hulst J. M., Sancisi R., 2005, A&A, 431, 65
 Bregman J.N., ApJ, 236, 577
 Collins J.A., Benjamin R.A., Rand R.J., 2002, ApJ, 578, 98
 Diehl R., et al., 2006, Nature, 439, 45
 Hernandez X., Valls-Gabaud D., Gilmore G., 2000, MNRAS, 316, 605
 Binney J., 2005, Phil. Trans. Roy. Soc. A, 363, 739
 Binney J., Tabor G., 1995, MNRAS, 276, 663
 Binney J., Tremaine S., 2008, *Galactic Dynamics*, Princeton University Press, Princeton
 Ciotti L., Ostriker J.P. 2011, arXiv1104.2238
 Fraternali F., Binney J., 2006, MNRAS, 366, 449
 Fraternali F., Binney J., 2008, MNRAS, 386, 935
 Gursky H., Kellogg E., Leong C., Tananbaum H., Giacconi, R., 1971, ApJL, 167, L65
 Kalberla P.M.W., Burton W.B., Hartmann D., Arnal E.M., Bajaja E., Morras R., Pöppel W.G.L., 2005, A&A, 440, 775
 Li Y.S., White S.D.M., 2008, MNRAS, 384, 1459
 Marasco A., Fraternali F., Binney J.J., 2011, MNRAS, in press
 Marinacci F., Binney, J., Fraternali F., Nipoti C., Ciotti L., Londrillo P., 2010, MNRAS, 404, 1464
 Marinacci F., Fraternali F., Nipoti C., Binney, J., Ciotti L., Londrillo P., 2010, MNRAS, 415, 1534
 Mastropietro C., Burkert A., Moore B., 2009, MNRAS, 399, 2004
 Omma H., Binney J., 2004, MNRAS, 350, L13
 Oosterloo T., Fraternali, F., Sancisi R., 2007, AJ, 134, 1019
 Pagel B., Patchett B.E., 1975, MNRAS, 172, 13
 Peacock J., 1999, *Cosmological Physics*, Cambridge University Press, Cambridge
 Peterson J.R., Kahn S.M., Paerels F.B.S., Kaastra J.S., Tamura T., Bleeker J.A.M., Ferrigno C., Jernigan J.G., 2003, ApJ, 590, 207
 Putman M.E., et al., 1998, Nature, 394, 6695
 Sancisi R., Fraternali F., Oosterloo T., van der Hulst T., 2008, A&AR, 15, 189
 Sembach K.R., Wakker B.P., Savage B.D., Richter P., Meade M., Shull J.M., Jenkins E.B., Sonneborn G., Moos H.W., 2003, ApJS, 146, 165
 Spitzer L., 1956, ApJ, 124, 20
 Sunyaev R.A., Zeldovich Ya.B., 1972, Comments on Astrophysics & Space Physics, 4, 173
 Westmeier T., Brüns C., Kerp J., 2005, A&A, 432, 937